心理学报 2021, Vol. 53, No. 1, 1−14
Acta Psychologica Sinica

https://dx.doi.org/10.3724/SP.J.1041.2021.00001

汉语口语词汇产生中的多重音韵激活：单词翻译任务的 ERP 研究*

张清芳[1] 钱宗愉[1] 朱雪冰[2]

([1] 中国人民大学心理学系，北京 100872) ([2] 上海外国语大学，上海 201620)

摘 要 口语词汇产生过程中，非目标词是否会产生音韵激活是独立两阶段模型和交互激活模型的争论焦点之一。研究运用事件相关电位技术，考察了被试在翻译命名任务中是否受到背景图片音韵或语义干扰词的影响。行为反应时中未发现显著的音韵效应，而语义效应显著，表明非目标词不会产生音韵激活。事件相关电位的结果显示在目标单词呈现后的 400~600 ms 时间窗口中出现了显著的语义效应，在 600~700 ms 时间窗口内出现了边缘显著的语义效应和音韵效应，均表现为相关条件比无关条件波幅更正。上述结果表明在将英语翻译成汉语的过程中，尽管在脑电上呈现出可能存在微弱的多重音韵激活，但行为结果并不会显示出非目标项的音韵激活。研究结果支持了汉语口语词汇产生遵循独立两阶段模式的观点。

关键词 口语词汇产生，多重音韵激活，独立两阶段模型，交互激活模型
分类号 B842

1 前言

口语产生是人们通过语音形式表达思想的过程，包括了概念化、形式化以及发音运动三个阶段。在概念化阶段，讲话者明确想要表达的概念；形式化阶段包括了词条选择，音韵编码以及语音编码。最后，通过声带运动将发音目标以声音的形式输出(Levelt et al., 1999; Roelofs, 1997)。研究者对于形式化阶段中词条选择和音韵编码之间的关系存在争论。词条选择是词汇的激活和选择，词汇信息包括了单词的意义和语法特征，音韵编码是将词汇转化为以音素或以音节为单元进行组织的顺序，这两个过程形成了口语词汇产生中的词汇通达过程，对两个过程之间的关系存在截然不同的两类观点(Dell, 1986, 1988; Levelt et al., 1999)。

交互激活模型认为在口语词汇产生中，词条选择和音韵编码两个过程之间存在交互作用，激活在各个表征水平之间的扩散是双向的，因此该模型预测存在目标项的语义激活和音韵激活，而且非目标项也能产生音韵上的激活，讲话者最终选择音韵激活程度最高的进行发音(Dell, 1986)。而独立两阶段理论认为这两个过程之间无交互作用，信息的激活在语义和音韵表征之间的联结是单向的，不存在从音韵表征到语义表征之间的激活反馈。对于目标词汇的选择发生在词汇选择阶段，选择之后仅仅针对目标词产生音韵激活并进行发音，非目标项不会产生音韵激活(Levelt et al., 1999)。本研究关注的问题是在汉语口语词汇产生过程中，非目标项是否会产生音韵上的激活。

1.1 印欧语系口语词汇产生中的多重音韵激活

已有研究表明印欧语系(例如英语、荷兰语、德语、西班牙语等)口语词汇产生过程中存在多重音韵激活(Jescheniak & Schriefers, 1998; Kuipers & La Heij, 2009; Navarrete & Costa, 2005; Peterson & Savoy, 1998)。研究者采用图画−词汇干扰范式(picture-word interference paradigm, PWI)、图−图干

收稿日期：2020-03-05
* 北京市社会科学基金重点项目(16YYA006)，中国人民大学科学研究基金项目(中央高校基本科研业务费专项) (18XNLG28)项目资助。
通信作者：张清芳，E-mail: qingfang.zhang@ruc.edu.cn；朱雪冰，E-mail: zhuxb@shisu.edu.cn

扰范式(picture-picture interference paradigm, PPI)等，设置与目标词的语义相关词存在音韵联系的单词作为干扰词条件，考察与无关条件相比是否会延长目标词的产生时间，来探测是否存在多重音韵激活。在这两类任务中，目标任务均为图画命名，唯一不同的是干扰项, PWI 中为词语，而 PPI 中为图画。

在 PWI 中，为了探测到非目标项的音韵激活，研究者所采取的图画一般都有两个名称，比如英语中的“sofa”和“couch”为同一幅图画的名称，这是为了最大程度地增加非目标项的语义激活，研究者认为语义激活程度高，其激活才更有可能扩散到音韵层面被探测到。采用 PWI, Jescheniak 和 Schriefers (1998)在德语的研究中用听觉范式呈现干扰词，发现在图画命名中非目标项的音韵干扰词显著地延长了图画命名的时间。为了确保图画的两个名称都被激活, Peterson 和 Savoy (1998)采用非同时呈现的 PWI，要求被试在图画命名任务和词汇命名任务间进行切换，实验中设置了两个音韵相关词，分别为“count”和“soda”，结果发现被试在图片呈现后命名单词“count”和“soda”时潜伏期都得到了缩短，表明与图画存在语义相关的名称“sofa”也产生了音韵激活。Kurtz 等 (2018)进一步强化被试对正确目标词的学习，即加强词条选择阶段目标词条的竞争优势后，在图画命名任务中探测到同义词的音韵激活。

研究者也采用 PPI 范式探测多重音韵激活。在 PPI 中，一般呈现两幅线条颜色不同的图片，两个图片名称存在音韵相关，要求被试对其中一种颜色的图片进行命名。Madebach 等(2011)操控了空间上分离的两张图片的可视度，发现命名不容易辨别的图片时，不能探测到非目标图片的音韵激活，这是由于加工非目标图片的认知资源较少，其语义激活较低，因此也不能探测到其音韵激活。Oppermann 等(2014)操控了目标图片与非目标图片外形的相似程度，发现给被试呈现两张外形相似的图片时(如“雨伞”和“棕榈树”)，探测到了非目标图片的音韵激活。研究者认为因为视觉上容易混淆，被试对于目标和非目标图片都分配了注意，使得非目标图片的激活增强，因而发现了非目标图片的音韵激活。Roelofs (2008)分析了被试命名图画时的眼动指标——凝视转移潜伏期(gaze shift latency)，当非目标图片与目标图片音韵相关时，缩短了命名潜伏期和凝视转移潜伏期，说明被试激活了非目标图片的音韵信息。上述研究表明注意资源的分配会影响对非目标图片的音韵激活。

研究者认为采用双语任务来考察非目标项的音韵激活更具有优势，双语者的第一与第二语言可以共享语义特征，从而可以使得非目标项的语义信息激活程度更高(Costa et al., 2000)。Costa 等的研究中要求卡特兰语−西班牙语(Catalan-Spanish)双语者和西班牙语单语者用西班牙语对同源词和非同源词进行命名，同源词指不同语言中语义相同，字形与音韵相似的词对。结果发现双语者在进行同源词条件的反应时显著小于非同源词条件，而单语者在同源与非同源词条件无差异。这表明在双语者的口语词汇产生过程中，非目标项产生了音韵激活：在同源词条件下，如命名目标项为西班牙语“gato”，其非目标语言卡特兰语中的“gat”也同时激活相应的音韵信息，从而对目标项的音韵信息(/g/, /a/, /t/)的提取产生促进作用，而非同源词条件中的非目标项与目标词没有任何共享的音素，因此未观察到启动作用，表明双语者可以对非目标项的音韵信息进行加工。Boukadi 等(2015)在突尼斯阿拉伯语−法语双语者中发现了相似的结果。

研究者采用单词翻译任务和单词联想任务来考察多重音韵激活，同时避免了 PPI 范式中同时呈现两张图片所带来的注意分配问题。采用单词翻译任务时，实验中呈现目标词和干扰词(或干扰图)，要求被试忽略干扰词(干扰图)对目标词进行翻译。目标词与干扰词(或干扰图的名称)之间为音韵相关和音韵无关。在荷兰语−英语双语情境下，研究者未发现背景图片的音韵激活(Bloem & La Heij, 2003; Bloem et al., 2004)。但是, Navarrete 和 Costa (2009)在西班牙语−卡特兰语双语语境下，采用相同的实验设计和任务，结果发现音韵相关的背景图片会加快翻译潜伏期，即背景图片的音韵得到激活。降低音韵相关比例后背景图片的音韵关系不再影响被试单词翻译潜伏期，但音韵无关条件较音韵相关条件被试的错误率显著增加。因此考察背景图片音韵能否激活采用单词翻译任务的印欧语系研究在反应时和错误率指标上探测到了非目标项的音韵激活。Humphreys 等(2010)采用词汇联想任务，呈现给被试一个英文单词，在其下方呈现干扰图片，要求被试进行单词联想任务，根据所呈现单词说出第一个联想到的单词，结果发现与无关条件相比，干扰图片与目标词音韵相关条件显著缩短了单词产生的潜伏期，表明英语口语词汇产生中非目标图片产生了音韵激活。

综上，印欧语系的大量研究都探测到了非目标

项的音韵激活，探测到这一激活的条件是要保证非目标图片产生较强的概念和语义水平上的激活。

1.2　汉语口语词汇产生中的多重音韵激活

汉语口语词汇产生中是否存在多重音韵激活，研究者得到的结果并不一致。庄捷和周晓林(2003)采用 PWI 任务发现语义中介的音韵相关条件(如目标图片为“牛”，干扰字为与目标字语义相关“羊”的同音字“阳”)与控制条件之间在反应时上的差异边缘显著。研究中也采用与 Peterson 和 Savoy (1998)类似的图画命名和汉字命名切换任务，结果发现语义中介音韵相关条件显著缩短了被试对汉字命名的潜伏期。庄捷和周晓林(2003)因此认为汉语口语词汇产生中存在非目标项的音韵激活。张清芳和杨玉芳(2006)采用PWI任务却未发现语义中介的音韵相关干扰字对图画命名潜伏期产生显著影响，但发现了与目标名称存在字形相关的干扰词产生了音韵激活，表明非目标项产生了音韵激活。

然而，近期的行为研究都未发现汉语口语词汇产生中非目标项的音韵激活。Zhang 和 Zhu (2016)考虑到采用 PWI 任务在汉语中很难找到同一幅图片对应两个名称的情况，以及 PPI 任务中图片注意分配的问题，采用了单词翻译任务和词汇联想任务均未发现汉语口语词汇产生过程中背景图片的多重音韵激活。Zhang 等(2018)为提高非目标项的语义激活水平，采用语义组块范式，设置了语义同质组，即在同一组块中出现的所有图片均属于相同的语义范畴，使词条选择阶段目标词与非目标词竞争加剧，更易出现多重音韵激活。为了探测到非目标项的音韵激活，Zhang 等结合了 PWI 任务，研究结果显示语义组块的设置提高了非目标项的语义激活水平，但仅在英语图画命名中发现了非目标项的音韵激活，汉语中却没有。

上述研究均采用行为反应时作为指标，事件相关电位(event-related potentials, ERP)技术具有高时间分辨率的特点，可以即时地反映认知加工过程中的变化，是更为敏感的测量指标。目前已有研究采用事件相关电位技术考察自然发音命名过程中词汇产生的具体时间进程(Costa et al., 2009; Dell’Acqua et al., 2010; Python et al., 2018; Zhu et al., 2015; 详见综述 Ganushchak et al., 2011)。这些研究主要考察了口语词汇产生中语义和音韵激活的时间进程，对于多重音韵激活的探索较少。仅有的一项 ERP 研究来自 Jescheniak 等(2003)，研究者采用延迟图画命名任务在德语中考察了非目标项(目标词语义相关词)是否会产生音韵激活，结果并未探测到。这可能是由于延迟图画命名任务不是一个自然的口语词汇产生过程，非目标项的音韵激活本身就比较微弱，影响了其激活强度。

目前尚未有研究采用 ERP 技术考察汉语口语词汇产生过程中的多重音韵激活。我们选择了单词翻译任务，给被试呈现一个英文单词和一幅图片，要求被试将英文单词翻译成汉语。在实验设置上，图和词二者容易分离，避免了 PPI 范式所带来的注意混淆或注意不足的问题。单词翻译过程包括词汇识别、(双语)概念激活、词条选择、音韵编码、语音编码及发音执行等过程。同时已有研究发现当非目标项以视觉形式呈现时更容易探测到音韵激活(Oppermann et al., 2008)，因此采用视觉呈现的单词翻译任务有利于探测到非目标项的多重音韵激活。研究者发现设置与目标词存在语义关系的图片会缩短单词翻译的潜伏期，表现出语义促进效应。这一效应是由于双语者的两类语言共享概念节点所产生的(Bloem & La Heij, 2003; Bloem et al., 2004; La Heij et al., 1996; Navarrete & Costa, 2009; Rahman & Melinger, 2009; Rahman & Melinger, 2019)。荷兰语−英语双语者中，Bloem 和 La Heiji (2003)在单词翻译任务中设置了两类线索，分别为线索图和线索词，被试的任务是把所呈现的英语单词翻译成荷兰语(母语)，结果发现在线索图片的条件下，出现了语义促进效应，而在线索词条件下，出现了语义抑制效应。这是因为线索图在概念水平上的激活促进了目标词的翻译过程，而线索词的加工是直接通达词条，导致词汇水平上的竞争，线索词的呈现对目标词的翻译产生了抑制效应，表明单词翻译任务下的语义促进效应发生在概念水平(同见 Navarrete & Costa, 2009)，而语义抑制效应发生在词汇水平(Roelofs, 1992; Starreveld & La Heiji, 1996)。

本实验主要目的是考察汉语口语词汇产生中非目标项是否产生音韵激活，在单词翻译任务中我们变化了英文单词对应的中文翻译词与图片名称之间的关系，包括了语义相关和无关，以及音韵相关和无关。设置语义相关和无关干扰图片的目的是为了重复已有的研究结果，验证被试是否注意到背景图片并进行加工，以保证实验设置的有效性。根据已有的行为研究结果(Zhang & Zhu, 2016)，我们预期会出现翻译潜伏期上的语义促进效应；若存在多重音韵激活现象，音韵相关条件会促进汉语对译词的产出，反应时变短，ERP 指标上也存在差异。

2 方法

2.1 被试

32 名来自北京地区高校的学生参与了本次实验(13 名男生, 19 名女生, 年龄 19~26 岁, 平均年龄 22.94 岁, 标准差 1.85 岁)。所有被试均为北方人, 普通话标准, 英语水平较高(英语六级成绩均高于 550 分), 视力或矫正视力正常, 实验后获得一定报酬。

2.2 材料

43 幅黑白线条图片选自张清芳和杨玉芳(2003)所建立的汉语标准图库, 其中 40 幅用于正式实验, 3 幅为练习图片, 图片名称均为双音节词。86 个英文单词选自 CELEX 数据库, 其中 40 个英文单词的汉语对译词与图片名称之间存在音韵相关(首字音节相同但声调不同)。例如, 图片名称为“飞机(/fei1ji1/)”, 英文单词为“soap”, 英文单词的汉语对译词为“肥皂(/fei2zao4/)”。音韵相关条件打乱后重新匹配形成音韵无关条件。另外 40 个英文单词汉语对译词与图片名称之间存在语义相关(二者具有语义联系, 属于相同语义范畴)。例如图片名称为“飞机(/fei1ji1/)”, 英文单词为“rocket”, 英文单词的汉语对译词为“火箭(/huo3jian4/)”, 二者均属于航天运输工具, 且不存在音韵或正字法上的关联。语义相关条件打乱后重新匹配形成语义无关条件。其余 6 个英文单词与 3 幅练习图片按照相同规则配对形成 12 对练习材料。实验材料的相关属性见表 1。

表 1 实验材料相关属性

实验材料属性	音韵条件	语义条件
英文单词平均 log 词频(次/百万)	1.62 ± 0.55	1.05 ± 0.57
英文单词音素数量	4.89 ± 1.61	4.44 ± 1.61
汉语对译词平均 log 词频(次/百万)	4.40 ± 0.59	3.70 ± 0.55
汉语对译词音节数	2	2

注：英文单词平均 log 词频查自 Neighborhood Watch program, Davis, 2005, 汉语对译词平均 log 词频查自《现代汉语通用词表》(2003)。

2.3 设计

实验为 2(相关类型：音韵、语义) × 2(相关性：相关、无关)的两因素完全重复测量设计, 两个自变量均为被试内因素。每组包括 12 个练习试次和 160 个实验试次。为了得到足够多的叠加次数, 每组重复 2 次, 两组之间有休息。刺激呈现的顺序是伪随机的：相同汉语目标词不会在连续 5 个试次中重复出现, 且声母相同的汉语目标词不连续出现。

2.4 仪器

实验程序由 E-Prime 1.1 编制, PST-SRBOX 反应盒, 麦克风和计算机。被试的反应通过 PST-SRBOX 连接的麦克风记录。实验材料的呈现、计时和被试反应时间均由计算机完成。主试记录被试反应的正确与否。

2.5 程序

被试距离电脑屏幕约 60 cm 完成实验任务。实验共分为三个阶段：(1)学习阶段：被试学习英文单词正确的汉语对译词和图画正确名称。(2)测试阶段：被试经过测试, 需要达到可以正确而迅速地翻译单词和命名图画。(3)正式实验阶段：首先屏幕正中间呈现注视点“+” 500 ms, 然后空屏 500 ms, 接着屏幕中央同时呈现一个英文单词和一幅图片。图片经过标准化形成大小为 6 cm × 6 cm 的图片, 英文单词字体为 Arial, 28 号字体, 呈现在图片中央的位置。被试的任务是忽略图片, 尽可能正确而迅速地对英文单词进行翻译, 被试做出反应时, 刺激消失。主试对被试的反应进行正误判断, 1500 ms 的间隔后进入下一试次。如果被试 5000 ms 内不做反应, 刺激自动消失, 主试按键后进入下一试次。实验持续时间约为 40 分钟。

2.6 EEG 记录与分析

使用Neuroscan公司生产的64导脑电记录系统, 银、氯化银电极以国际通用的 10-20 方式固定于电极帽上。左侧乳突为参考电极, 额头中央连接接地电极。位于左眼上下眼眶正中间的电极记录垂直眼电(VEOG), 位于左右眼角两侧外 1 cm 处的电极记录水平眼电(HEOG)。每个电极与头皮之间的阻抗小于 5 kΩ。连续记录时滤波带通为 0.05~70 Hz, 采样率为 500 Hz。

采用 EEGLAB 软件对 EEG 数据进行离线分析, 选择双侧乳突电极进行重参考。通过 EEGLAB 软件进行眼电伪迹矫正, 滤波带通为 0.1~30 Hz, 按照刺激呈现前 200 ms 和呈现后 1000 ms 的时间段对脑电进行分段, 前 200 ms 作为基线进行校正, 删除眼电、肌电、漂移等伪迹, 删除波幅超过±100 μV 的试次, 仅对被试正确反应试次的 EEG 数据进行平均叠加。

选取脑区(前、中、后)和半球(左、中、右) 9 个兴趣区进行分析, 其中每个外侧兴趣区的波幅均为邻近的 3 个电极平均值：左前(F3, F5, FC3), 中前(Fz), 右前(F4, F6, FC4), 左中(C3, C5, CP3), 中

中(Cz), 右中(C4, C6, CP4), 左后(P3, P5, PO3), 中后(Pz), 右后(P4, P6, PO4)。以各时间窗口的平均波幅为因变量, 进行 2(相关类型: 音韵、语义) × 2(相关性: 相关、无关) × 3(脑区: 前、中、后) × 3(半球: 左、中、右)的重复测量方差分析。当重复测量方差分析中球形性假设不成立时, 采用 Greenhouse-Geisser 校正, 并采用 R 软件中的 fdrtool 安装包对多重比较中的 p 值进行校正。

3　结果

32 名被试中删除了 5 名 EEG 信号伪迹过多的被试数据, 共分析了 27 名被试的数据。每名被试每种条件下有效叠加试次均超过 40 次。

3.1　行为结果

删除错误反应的数据 1.24% (4 种条件下错误率分别为: 音韵相关 1.30%, 音韵无关 1.34%, 语义相关 0.93%, 语义无关 1.39%), 删除反应时小于 200 ms 和大于 2000 ms 的数据 1.02%, 删除平均值三个标准差以外的数据 1.33%。

为同时考虑被试和实验材料所产生的随机效应, 我们使用 R 软件中的 lme4 程序包, 对反应时数据进行混合线性模型拟合分析(Baayen et al., 2008)。模型拟合主要包括三个步骤: 首先, 指定一个只包含随机因素(被试和项目)的零模型; 第二, 通过添加固定因子来丰富零模型。在已有模型的基础上, 逐步增加两个自变量及其交互作用(相关类型; 相关性; 二者交互作用)。第三, 使用卡方检验将新建立的模型与之前的模型进行比较。若在已有模型中加入固定因子或两个因素的交互作用对方差估计的改善不显著, 则当前模型是最佳拟合模型。

反应时分析结果显示, 在相关类型和相关性进入模型的基础上加入二者交互作用, 可以显著改变模型拟合程度, $\chi^2(1)= 13.07$, $p < 0.001$。最终确定的最优拟合模型[RT~(相关类型+相关性)^2 + (1|被试) + (1|项目)]中包含了变量相关类型和相关性, 以及两者的交互作用。相关类型的主效应显著(β = 19.07, t = 3.47, p < 0.001), 相关类型和相关性的交互作用显著(β = 28.17, t = 3.62, p < 0.001), 相关性主效应不显著(β = 3.49, t = 0.64, p = 0.525)。进一步比较发现, 相关类型为音韵条件下, 相关与无关条件之间不存在显著差异(917 − 921= −4 ms, β = 3.23, t = 0.62, p = 0.534), 相关类型为语义条件下, 相关与无关条件之间存在显著差异(935 − 967= −32 ms, β = 31.75, t = 5.70, p < 0.001), 结果见图 1。

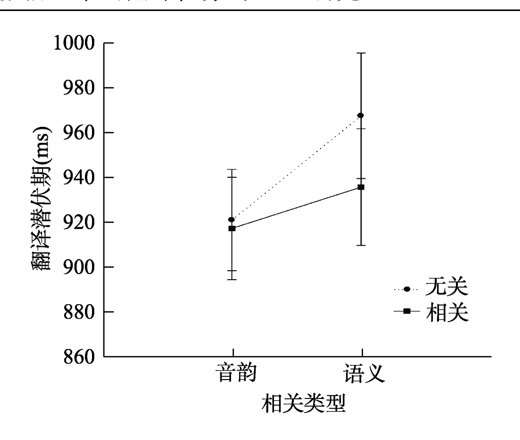


图 1　不同条件下单词翻译潜伏期
注: 图中误差棒为 95% CI

为排除重复次数对我们所关注音韵和语义效应产生潜在的影响, 我们从只包含随机因子(被试和项目)的零模型起步, 逐步增加固定因子(相关类型; 相关性; 重复次数; 二重和三重交互作用)。在三个自变量和二重交互作用进入模型的基础上加入三重交互作用, 可以显著改变模型拟合程度, $\chi^2(1)$ = 6.79, p = 0.009。最终确定的最优拟合模型[RT~(相关类型+相关性+重复次数)^3 + (1|被试) + (1|项目)]包括了相关类型、相关性、重复次数、相关类型和相关性的交互作用、相关类型和重复次数的交互作用、相关性和重复次数的交互作用和三个变量之间的交互作用(统计结果见表 2)。针对显著的三重交互作用, 进一步比较发现: 当相关类型为音韵条件下, 第一次呈现相关与无关条件之间的单词翻译潜伏期不存在显著差异(β = −1.14, t = −0.15, p = 0.884), 第二次呈现相关与无关条件之间的单词翻译潜伏期不存在显著差异(β = 8.29, t = 1.32, p = 0.189); 当相关类型为语义条件下, 第一次呈现单词翻译潜伏期相关与无关条件之间存在显著差异(β = 47.02, t = 5.49, p < 0.001), 第二次呈现相关与无关条件之间的单词翻译潜伏期存在显著差异(β = 16.06, t = 2.46, p = 0.014)。由此可见, 重复次数不影响是否存在音韵效应或语义效应。

针对不显著的音韵效应, 进一步使用 JASP 软件(https://jasp-stats.org/, JASP Team 2017) (JASP Team, 2017; Marsman & Wagenmakers, 2017; Wagenmakers et al., 2018)进行贝叶斯配对样本 t 检验, 以计算对零假设的支持程度。根据 Jeffreys (1961)提出的贝叶斯因子分类标准认为 BF_{01} 的值位于 3~10 之间则表明存在实质性证据支持零假设。本研究中统计结果显示 BF_{01}= 3.54, 表示零假设(假定不存在音韵效应的)成立的可能性是备择假设(假

表 2　以单词翻译潜伏期为因变量的混合线性模型的固定效应(考虑重复次数时)

自变量	β	SE	*df*	*t*	*p*
截距	955.72	14.67	52.35	65.13	< 0.001
相关类型	13.40	7.61	8254.62	1.76	0.078
相关性	−0.85	7.59	8254.33	−0.11	0.911
重复次数	−77.95	7.56	8254.10	−10.31	< 0.001
相关类型×相关性	47.64	10.79	8254.57	4.42	< 0.001
相关类型×重复次数	12.10	10.72	8254.17	1.13	0.259
相关性×重复次数	9.03	10.70	8254.12	0.84	0.399
相关类型×相关性×重复次数	−39.58	15.20	8254.18	−2.60	0.009

定存在音韵效应的)成立可能性的 3.54 倍，存在实质性证据支持零假设，表明背景图片与目标词音韵相关或无关对翻译潜伏期的影响不存在差异。

由于错误率过低，未对错误率做进一步分析。

3.2　ERP 结果

剔除反应错误和反应时小于 500 ms 或大于 2000 ms 的试次，以及反应时在三个标准差以外的试次。由于口语产生涉及肺、喉、声带等肌肉和发音器协同运动，发音所产生的运动伪迹会影响靠近发音的 EEG 信号(Ouyang et al., 2016)，我们选定 ERP 波幅分析的截止时间为刺激呈现后的 700 ms，以 100 ms 为间隔分别分析刺激呈现后 0~700 ms 内的 ERP 波幅，共确定 7 个时间窗口。对每个时间窗口的平均波幅进行 2(相关类型：音韵、语义) × 2(相关性：相关、无关) × 3(脑区：前、中、后) × 3(半球：左、中、右)的重复测量方差分析，结果如下：

在 0~100 ms 时间窗口内，相关类型、相关性和脑区三者交互作用显著 $F(2, 52) = 4.67$, $p = 0.033$, $\eta_p^2 = 0.152$；相关类型、相关性和半球三者交互作用显著 $F(2, 52) = 3.29$, $p = 0.045$, $\eta_p^2 = 0.112$。在 100~200 ms 时间窗口内，相关类型主效应边缘显著，$F(1, 26) = 3.60$, $p = 0.069$, $\eta_p^2 = 0.122$；相关类型和相关性交互作用边缘显著，$F(1, 26) = 3.65$, $p = 0.067$, $\eta_p^2 = 0.123$；相关类型和半球交互作用显著，$F(2, 52) = 9.29$, $p < 0.001$, $\eta_p^2 = 0.263$。在 200~300 ms 时间窗口内，相关类型主效应显著，$F(1, 26) = 6.82$, $p = 0.015$, $\eta_p^2 = 0.208$；相关类型和半球交互作用显著，$F(2, 52) = 7.00$, $p = 0.004$, $\eta_p^2 = 0.212$。在 300~400 ms 时间窗口内，相关性主效应边缘显著，$F(1, 26) = 2.97$, $p = 0.097$, $\eta_p^2 = 0.103$；相关类型和半球的交互作用边缘显著，$F(2, 52) = 3.11$, $p = 0.066$, $\eta_p^2 = 0.107$。在 400~500 ms 时间窗口内，相关性主效应显著，$F(1, 26) = 11.62$, $p = 0.002$, $\eta_p^2 = 0.309$；相关类型和相关性交互作用边缘显著，$F(1, 26) = 3.02$, $p = 0.094$, $\eta_p^2 = 0.104$。在 500~600 ms 时间窗口内，相关性主效应显著，$F(1, 26) = 9.94$, $p = 0.004$, $\eta_p^2 = 0.277$；相关类型和相关性交互作用显著，$F(1, 26) = 6.75$, $p = 0.015$, $\eta_p^2 = 0.206$；相关性和半球交互作用显著 $F(2, 52) = 3.22$, $p = 0.048$, $\eta_p^2 = 0.110$；相关类型、相关性和脑区三者交互作用边缘显著，$F(2, 52) = 3.31$, $p = 0.070$, $\eta_p^2 = 0.113$。在 600~700 ms 时间窗口内，相关性主效应显著，$F(1, 26) = 6.26$, $p = 0.019$, $\eta_p^2 = 0.194$。

根据存在的交互作用，在 7 个时间窗口中分别做简单效应分析，计算各个兴趣区内的音韵效应和语义效应，结果显示在 400~600 ms 时间窗口内出现显著的语义效应，600~700 ms 时间窗口内出现边缘显著的音韵效应和语义效应(见表 3)。针对各个时间窗口内显著和边缘显著的音韵效应和语义效应，进一步使用 JASP 软件进行贝叶斯配对样本 *t* 检验，计算对备择假设的支持程度(见表 4)。结果显示，处于边缘显著水平的音韵效应对应的 BF_{10} 最小值为中中区域的 2.01，最大值为右中区的 3.27，表明存在比较可靠的证据支持了音韵效应在右中区的存在。处于显著和边缘显著水平的语义效应对应的 BF_{10} 值最小为左前区域的 0.93，最大为 500~600 ms 时间窗口内中中区域的 50.25，多个区域包括左中、左后、中前、中中以及右中区域的 BF_{10} 值在 400~600 ms 时间窗口内都大于 10，表明有强证据支持了语义效应的存在(Jeffreys, 1961, P.432)。上述结果表明，尽管在神经层面上在某个特定兴趣区保留了多重音韵激活，但在行为结果上不足以产生多重音韵激活的效应，相比而言，语义效应的脑电激活模式分布范围广且效应强，在行为结果上也表现为显著的促进效应。

图 2 所示分别为语义相关和语义无关条件，音韵相关和音韵无关条件在三个兴趣区内的 ERP 波形，以及语义效应、音韵效应差异波(相关−无关)

表 3　400~700 ms 时间窗口内不同兴趣区的音韵效应和语义效应

时间窗口	ROIs																	
	左						中						右					
	前		中		后		前		中		后		前		中		后	
	F	η_p^2	F	η_p^2	F	η_p^2	F	η_p^2	F	η_p^2	F	η_p^2	F	η_p^2	F	η_p^2	F	η_p^2
音韵效应																		
400~500 ms	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
500~600 ms	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
600~700 ms	—	—	—	—	—	—	—	—	$5.45^{\dagger}$	0.17	—	—	$5.65^{\dagger}$	0.18	$6.75^{\dagger}$	0.21	—	—
语义效应																		
400~500 ms	$3.96^{\dagger}$	0.13	12.67^{**}	0.33	12.99^{**}	0.33	11.56^{**}	0.31	10.00^{**}	0.28	4.67^{*}	0.15	6.55^{*}	0.20	7.07^{*}	0.21	5.97^{*}	0.19
500~600 ms	$3.50^{\dagger}$	0.12	14.20^{**}	0.35	11.71^{**}	0.31	8.34^{*}	0.24	15.03^{**}	0.37	8.35^{*}	0.24	6.66^{*}	0.20	11.72^{**}	0.31	6.81^{*}	0.21
600~700 ms	—	—	$5.81^{\dagger}$	0.18	$5.90^{\dagger}$	0.19	—	—	—	—	—	—	—	—	$5.48^{\dagger}$	0.17	—	—

注：$^{**}p < 0.01$, $^{*}p < 0.05$, $0.05 < {}^{\dagger}p < 0.10$, —不显著，双侧检验，$p$ 值经 FDR 校正。

表 4　400~700 ms 时间窗口内不同兴趣区音韵效应和语义效应的 BF_{10} 值

时间窗口	ROIs								
	左			中			右		
	前	中	后	前	中	后	前	中	后
音韵效应									
400~500 ms	—	—	—	—	—	—	—	—	—
500~600 ms	—	—	—	—	—	—	—	—	—
600~700 ms	—	—	—	—	2.01	—	2.17	3.27	—
语义效应									
400~500 ms	1.12	24.44	27.00	17.18	10.29	1.49	3.04	3.68	2.45
500~600 ms	0.93	39.14	18.03	5.82	50.25	5.83	3.17	18.09	3.35
600~700 ms	—	2.30	2.38	—	—	—	—	2.04	—

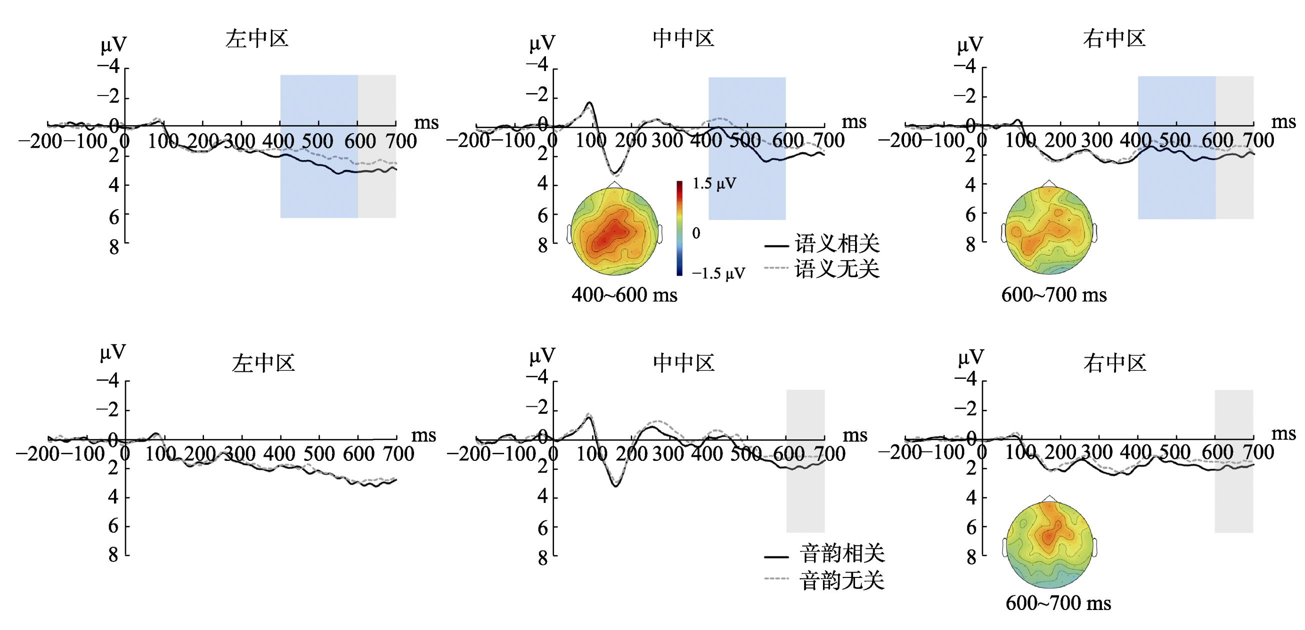


图 2　语义效应和音韵效应的 ERP 波形和差异波地形图

注：蓝色阴影部分为 $p < 0.05$，灰色阴影部分为 $0.05 < p < 0.1$, p 值经 FDR 校正

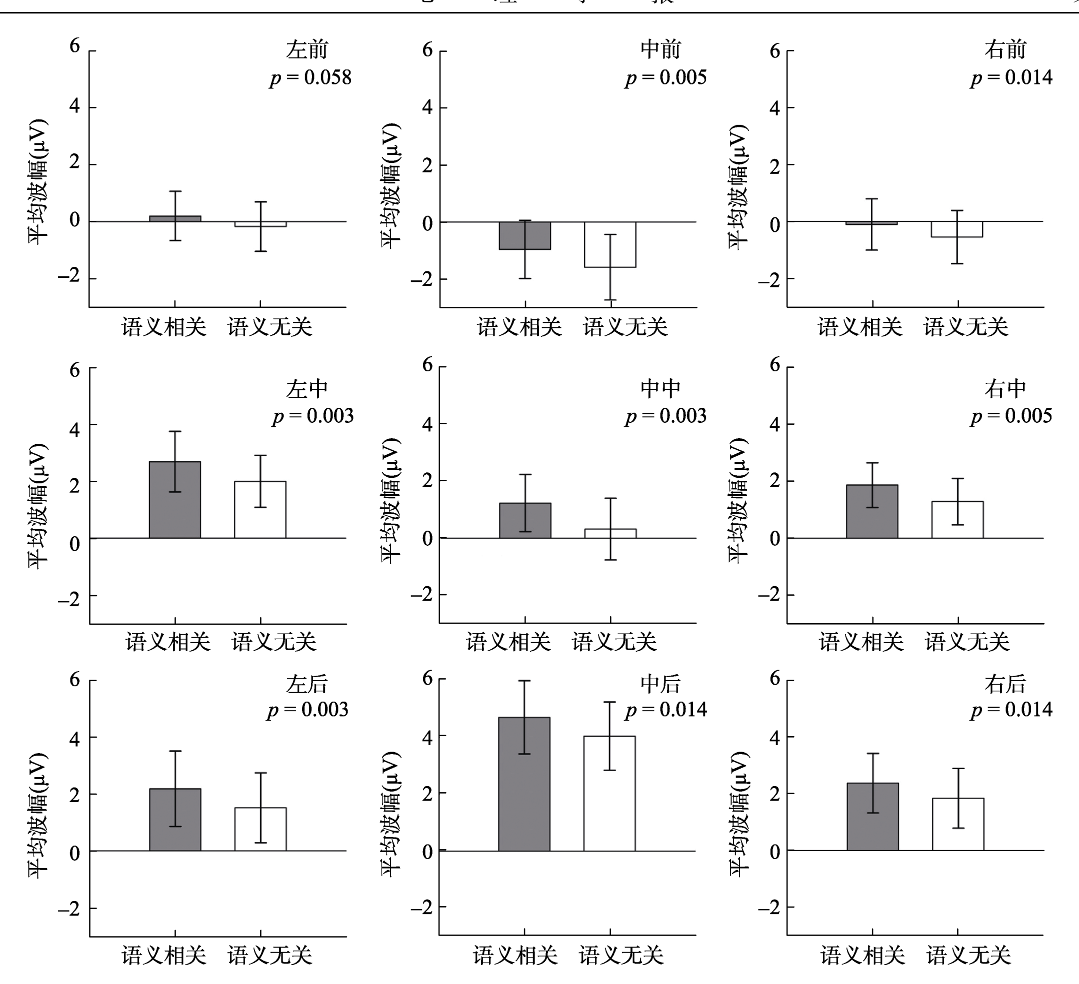


图 3　400~600 ms 时间窗口下不同兴趣区语义效应的平均波幅

注：图中误差棒均为 95% CI, p 值经 FDR 校正

地形分布图。图 3 所示为 400~600 ms 时间窗口下各个兴趣区达到显著水平的语义效应。

3.3　反应时和 ERP 波幅的相关

显著的语义效应出现在目标刺激呈现后的 400~600 ms 时间窗口，因此我们对这一时间窗口内显著的语义效应进行反应时和 ERP 的相关分析，以超过 3 个标准化残差为标准剔除奇异值，结果发现，中后区语义相关与语义无关条件下的反应时差值和对应的 ERP 差异波之间存在显著的负相关(r (26) = −0.52, p = 0.006, 95% CI = [−0.77, −0.25]), FDR 校正后边缘显著，p = 0.054，语义相关与无关反应时差值越负，ERP 差异波越正(见图 4)，表明反应时越短，波形越正。其余兴趣区内无显著的相关关系(r < |0.29|, ps > 0.15, FDR 校正后 ps > 0.681)。

4　讨论

本研究采用单词翻译任务，利用 ERP 技术考察了双语情境下汉语词汇产生过程中是否存在多重音韵激活。行为结果表明存在语义促进效应，但

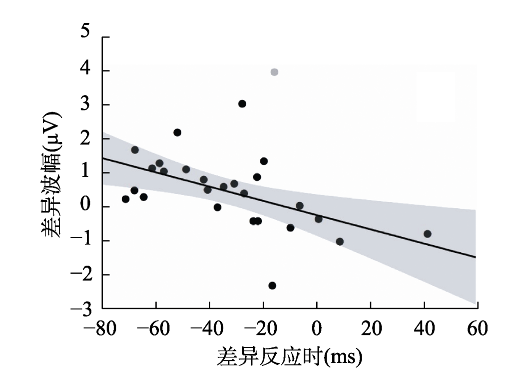


图 4　400~600 ms 时间窗口内在中后区域语义相关与语义无关条件间的翻译潜伏期差值与对应条件的 ERP 波幅差值之间的相关。灰色散点为剔除的奇异值。

是不存在多重音韵激活。ERP 结果表明语义相关和语义无关条件间的差异发生在英文单词呈现后的 400~600 ms 时间窗口内，表现为语义相关条件比语义无关条件下平均波幅更正；音韵相关条件和音韵无关条件下的波形在 600~700 ms 窗口内出现了边缘显著的差异，表现为音韵相关条件比音韵无关

条件下平均波幅更正，在脑电上讲话者可能保留了多重音韵激活的模式。

4.1　多重音韵激活

在单词翻译任务中，行为结果表明不存在非目标项的音韵激活，这与已有汉语的研究结果一致(Zhang & Zhu, 2016; Zhang et al., 2018)，支持了独立两阶段理论。独立两阶段理论假设词汇选择和音韵编码之间的激活传递是单向的，目标词的选择是在词汇选择阶段完成的，仅仅针对目标项产生音韵激活。已有的研究结果表明汉语的口语词汇产生过程遵循独立两阶段的模式，这表现在两个方面：第一，本研究和已有研究(张清芳，杨玉芳，2006; Zhang & Zhu, 2016; Zhang et al., 2018)在行为指标上均未发现非目标项的多重音韵激活。第二，词汇选择阶段的语义激活和音韵编码阶段的音韵激活在时间上不存在重叠(Zhu et al., 2015; Zhu et al., 2016)。行为和 ERP 的研究都发现在汉语口语词汇产生中，语义效应和音韵效应是独立的，两者之间不存在交互作用，两个效应分别出现于图画呈现后 250~450 ms 和 450~600 ms，时间窗口上不存在重叠，符合独立两阶段模型的预期。

本研究结果与庄捷和周晓林(2003)的发现不同。庄捷等的研究中音韵相关条件下语义相关项和干扰字的发音是完全一样的，例如语义相关项为“羊”，干扰字为“阳”，二者音节音调完全相同，共用一个音韵表征，这给被试提供了较强线索通过语义激活去影响图画命名过程，即语义中介词的干扰效应是通过干扰词的音韵激活，进而激活了相关的语义。也就是说，干扰字“阳”的音韵激活传递到“羊”的语义表征上，增加了“羊”的语义激活，与无关条件下相比，这一条件与目标词“牛”的激活形成了更强的竞争关系，因而产生了抑制效应。由于这一可能性，这一研究并不能明确地表明汉语口语词汇产生中存在非目标项的音韵激活。

与行为结果不一致的是，采用 ERP 技术探测到在大脑右中区域的电极上出现了较为可靠的证据表明了多重音韵激活的存在，表明讲话者可能在脑电上保留了多重音韵激活，但在行为反应时上并不会表现出来。这种行为和脑电之间的不一致与一些已有口语词汇研究的发现类似(Cai et al., 2020; Qu et al., 2012; Zhang & Damian, 2019)，我们认为这是由于脑电指标反映了在线的实时加工，而反应时是最终输出的结果反映，会受到一系列加工过程的影响。基于已有研究的结果和本研究的分析，我们倾向于认为讲话者在汉语口语词汇产生过程中，其神经层次上仍然保留了对多重音韵激活的敏感性。

使用类似的任务不能探测到汉语口语词汇产生中非目标项的音韵激活，但能探测到印欧语系中的这一激活，这表明口语词汇产生中非目标项的音韵激活在汉语中比较微弱，而在印欧语系中较强。这一结果模式与已有发现一致。汉语口语词汇产生的研究表明词汇选择和音韵编码两个阶段之间无交互作用，语义激活和音韵激活之间无时间上的重叠(Zhu et al., 2015; Zhu et al., 2016)。与汉语不同，印欧语系的语言中发现这两个阶段之间存在交互作用(Damian & Martin, 1999; Starreveld & La Heij, 1995, 1996)，且语义激活和音韵激活几乎是同时产生的(Dell’Acqua et al., 2010)。

汉语和印欧语系的口语产生分别遵循了独立两阶段模式和交互激活的模式，这主要涉及到语言加工过程中语义和音韵之间的联结强度。在汉语口语词汇产生过程中，语义和音韵之间的联结强度可能较弱，而在印欧语系中语义和音韵之间的联结可能较强。词汇产生与词汇理解过程都涉及到在心理词典中提取相应的正字法、语义和音韵这三类信息之间的联结及其加工。例如，在图画−词汇干扰实验任务中，采用图画命名任务，设置与目标字存在语义相关、正字法相关、音韵相关的干扰字以及无关字，以视觉方式呈现干扰字，结果发现了在图画命名这一经典的口语词汇产生任务中出现了显著的语义抑制效应、正字法促进效应和音韵促进效应，表明口语词汇产生过程中讲话者激活了目标词的字形信息，且促进了图画命名(汉语：Zhang et al., 2009; Zhang & Weekes, 2009; Zhao et al., 2012; 英语：Damian & Bowers, 2003)。在书写产生过程中，尽管最终输出的是字形信息，研究者同样发现了书写产生过程中音韵信息的激活(汉语：Qu et al., 2011; Wang & Zhang, 2015; 英语：Zhang & Damian, 2010)。

不同模式与不同语言的特点可能有密切联系。汉字是表意文字，用形旁表意是其显著的特征(Zhou & Marslen-Wilson, 1999)，尽管汉字的声旁也能在一定程度上提示语音，但是汉字字形和语音之间并未形成紧密的对应关系(Zhou & Marslen-Wilson, 2000)。例如，正字法类似的字，其发音可以完全不同(例如“床”和“庆”)；正字法完全不同的字，发音却可以相似(例如“床”和“闯”)。由于这样的特点，使得词汇的语义和语音之间的联系比较松散。相比而言，拼音文字由于存在形音对应的规则，形

音的对应关系十分密切(陈宝国 等, 2003), 这使得语义和语音之间的联结相对紧密。这些特点不仅会影响词汇产生过程(Zhu, et al., 2015, 2016), 而且对词汇阅读过程产生影响(陈宝国 等, 2003)。在语言加工过程中, 人们经过多次地在心理词典中激活语义和语音, 形成了语言加工的联结网络, 呈现出不同的联结强度。上述语言特点的不同导致了人们在使用语言的过程中语义和音韵之间的联结强度不同, 激活的扩散程度也不同, 导致了不同的口语词汇产生模式。

本研究使用单词翻译任务考察汉语口语词汇产生过程, 这一任务的优点是能保证被试反应的一致性。与此同时, 这使得在词条选择阶段参与竞争的非目标词数量较少, 被试能快速选出目标词(Jost et al., 2018), 研究者认为词条选择阶段时间过短很可能会难以探测到微弱的多重音韵激活现象(Jescheniak et al., 2006)。此外, 研究中所使用背景图片的名称均为双字词, 早期发现了汉语中存在多重音韵激活的研究中(庄捷, 周晓林, 2003)作为参与竞争的非目标词为单字, 单字和干扰字之间的音韵重叠程度显然高于双字词中仅有第一个字重叠的情况, 导致本研究中非目标项的音韵激活减弱。

4.2 语义促进效应

我们使用单词翻译任务发现了语义促进效应, 表现为语义相关条件下单词翻译潜伏期显著短于语义无关条件, 这一发现与已有研究结果一致(Bloem & La Heij, 2003; Nararrete & Costa, 2009)。汉英双语者在英语翻译成汉语的过程中出现与其他双语者类似的语义促进效应, 这一效应是通过双语者的英语和汉语所共享的概念层产生的(Christoffels et al., 2013; La Heij et al., 1996)。ERP 的结果发现这一过程发生在单词呈现后的 400~600 ms 之间, 这与采用图画命名任务的口语词汇产生过程不同。在单词翻译任务中, 目标项是词语, 被试需要首先再认出词语, 然后再激活相应的概念和对应的翻译词, 所以概念层的激活较晚。

已有采用翻译任务的 ERP 研究多探测的是词汇理解过程, 较少有研究采用词语翻译任务考察口语词汇产生过程。Christoffels 等(2013)首次采用 ERP 技术考察荷兰语–英语双语者单词翻译的时间进程, 研究中操纵了词汇的同形异义条件(如"room"在英语中翻译为"房间", 在荷兰语中翻译为"冰淇淋")和翻译方向(一语翻译成二语, 二语翻译成一语)两个变量, 结果发现同形异义词相较于控制组(异形异义词)的反应时更长, N400 波幅更负, 研究者认为是由于同形异义条件下词汇–语义阶段竞争更为激烈, 因而相比异形异义词产生的 ERP 波形更大, 而本研究中发现的是语义相关条件下 400~600 ms 时间窗口内的波幅更正。两个研究相比表明本研究中的语义促进效应发生在概念水平, 而非词汇–语义竞争阶段。尽管如此, Jost 等(2018)的研究中比较了翻译任务和单语产生任务发现两种任务在 424~630 ms 时间窗口下存在差异, 他们以图画命名中音韵编码阶段的时间进程为参考(Indefrey & Levelt, 2004), 认为单词翻译任务中 424~630 ms 时间窗口下的音韵编码阶段需要更多的自我监控。这一研究中同时比较了一语翻译成二语与二语翻译成一语两种翻译任务, 发现两种条件下的 ERP 波形在 600 ms 左右出现了差异, 研究者认为这反映了两种不同的翻译任务下概念水平的激活参与的程度, 这一解释与我们的研究发现一致。我们认为翻译任务中 424~630 ms 时间内的波形并非反映了对音韵编码阶段的监测。如果是的话, 概念加工阶段应该早于这一时间窗口, 而非 600 ms 左右。目前有关单词翻译任务的时间进程研究较少, 其加工过程与单语产生过程不同, 与单语词汇产生任务的时间进程做对比并不合适, 有关单词翻译过程中各个认知加工过程的时间进程需要进一步的研究。

在反应时和 ERP 波幅的相关分析中, 我们发现中后区显著的负相关(p = 0.006)经 FDR 校正后边缘显著(p = 0.054), 但在中后区还是可以观测到反应时和 ERP 波幅之间的线性趋势, 行为上语义促进效应越强, 400~600 ms 时间窗口内的差异波幅越大。根据图 2 中语义效应的地形图, 中中、左中、左后三个兴趣区内差异波幅最大, 但与反应潜伏期的相关分析却没有达到显著。虽然本研究未发现反应时和 ERP 波幅存在显著相关, 但结合单词翻译任务中概念激活的时间窗口应晚于图画命名任务下概念激活的时间窗口, 如 Bloem 和 La Heij (2003)的实验 3 通过操控目标刺激与干扰刺激呈现的时间间隔(stimulus onset asynchrony, SOA)发现单词翻译任务中概念激活的速度比图片概念激活的速度慢, 所以我们倾向于认为语义相关和语义无关出现差异的 400~600 ms 时间窗口为单词翻译过程中概念的激活。口语词汇产生的 ERP 研究很少, 反应时和波幅之间的相关分析结果需要更进一步的研究验证。

综上, 单词翻译任务中首先是单词识别过程,

之后是概念准备、词条选择和音韵编码阶段(见图 5)。概念准备阶段发生在词汇呈现后的 400~600 ms 之间，表明背景图的语义产生了激活，产生了语义促进效应。同时，在英语翻译成汉语的过程中，尽管在脑电上呈现出可能存在微弱的多重音韵激活，但行为结果并未显示出非目标项的音韵激活。研究结果支持了汉语口语词汇产生遵循独立两阶段模式的观点。

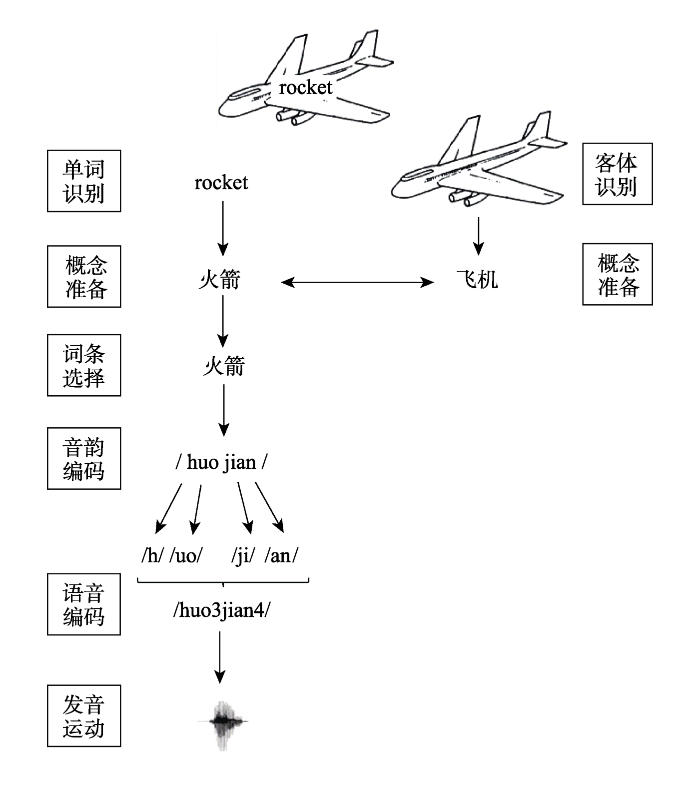


图 5　图词干扰范式下单词翻译任务的汉语口语词汇产生模型

参 考 文 献

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, *48*(3), 468–488.

Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. *Journal of Memory and Language*, *51*(2), 307–323.

Boukadi, M., Davies, R. A. I., & Wilson, M. A. (2015). Bilingual lexical selection as a dynamic process: Evidence from Arabic-French bilinguals. *Canadian Journal of Experimental Psychology*, *69*(4), 297–313.

Cai, X., Yin, Y. L., & Zhang, Q. F. (2020). The roles of syllables and phonemes during phonological encoding in Chinese spoken word production: A topographic ERP study. *Neuropsychologia*, *140*, 107382

Chen, B. G., Wang, L. X., & Peng, D. L. (2003). The time course of graphic, phonological, and semantic information processing in Chinese character recognition (II). *Acta Psychologica Sinica, 35(5),* 576–581.

[陈宝国, 王立新, 彭聃龄. (2003). 汉字识别中形音义激活时间进程的研究(II). *心理学报*, *35*(5), 576–581.]

Chinese Linguistic Data Consortium. (2003). *现代汉语通用词表[Chinese lexicon] (CLDC-LAC-2003–001)*. Beijing, China: Tsinghua University, State Key Laboratory of Intelligent Technology and Systems, and Chinese Academy of Sciences, Institute of Automation.

Christoffels, I. K., Ganushchak, L., & Koester, D. (2013). Language conflict in translation: An ERP study of translation production. *Journal of Cognitive Psychology*, *25*(5), 646–664.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology-Learning Memory and Cognition*, *26*(5), 1283–1296.

Costa, A., Strijkers, K., Martin, C. D., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(50), 21442–21446.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, *49*(1), 119–132.

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*(2), 345–361.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*, 65–70.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321.

Dell, G. S. (1988). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. *Journal of Memory and Language*, *27*(2), 124–142.

Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., & Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture-word interference paradigm. *Frontiers in Psychology*, *1*, 177

Ganushchak, L. Y., Christoffels, I. K., & Schiller, N. O. (2011). The use of electroencephalography in language production research: A review. *Frontiers in Psychology*, *2*, 208.

Humphreys, K. R., Boyd, C. H., & Watter, S. (2010). Phonological facilitation from pictures in a word association task: Evidence for routine cascaded processing in spoken word production. *Quarterly Journal of Experimental Psychology*, *63*(12), 2289–2296.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101–144.

JASP Team. (2017). JASP (Version 0.11.1) [Computer software].

Jeffreys, H. (1961). *Theory of probability*. Oxford, UK: Oxford University Press.

Jescheniak, J. D., Hahne, A., Hoffmann, S., & Wagner, V. (2006). Phonological activation of category coordinates during speech planning is observable in children but not in adults: Evidence for cascaded processing. *Journal of Experimental Psychology-Learning Memory and Cognition*,

32(2), 373–386.

Jescheniak, J. D., Hahne, A., & Schriefers, H. (2003). Information flow in the mental lexicon during speech planning: Evidence from event-related brain potentials. *Cognitive Brain Research, 15*(3), 261–276.

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology-Learning Memory and Cognition, 24*(5), 1256–1274.

Jost, L. B., Radman, N., Buetler, K. A., & Annoni, J. M. (2018). Behavioral and electrophysiological signatures of word translation processes. *Neuropsychologia, 109*, 245–254.

Kuipers, J. R., & La Heij, W. (2009). The limitations of cascading in the speech production system. *Language and Cognitive Processes*, *24*(1), 120–135.

Kurtz, F., Schriefers, H., Madebach, A., & Jescheniak, J. D. (2018). Incremental learning in word production: Tracing the fate of non-selected alternative picture names. *Journal of Experimental Psychology: Human Perception and Performance, 44*(10), 1586–1602.

La Heij, W., Hooglander, A., Kerling, R., & van der Velden, E. (1996). Nonverbal context effects in forward and backward word translation: Evidence for concept mediation. *Journal of Memory and Language, 35*(5), 648–665.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1–75.

Madebach, A., Jescheniak, J. D., Oppermann, F., & Schriefers, H. (2011). Ease of processing constrains the activation flow in the conceptual-lexical system during speech planning. *Journal of Experimental Psychology-Learning Memory and Cognition, 37*(3), 649–660.

Marsman, M., & Wagenmakers, E-J. (2017). Bayesian benefits with JASP. *European Journal of Developmental Psychology, 14*(5), 545–555.

Navarrete, E., & Costa, A. (2005). Phonological activation of ignored pictures: Further evidence for a cascade model of lexical access. *Journal of Memory and Language, 53*(3), 359–377.

Navarrete, E., & Costa, A. (2009). The distractor picture paradox in speech production: Evidence from the word translation task. *Journal of Psycholinguistic Research, 38*(6), 527–547.

Oppermann, F., Jescheniak, J. D., & Gorges, F. (2014). Resolving competition when naming an object in a multiple-object display. *Psychonomic Bulletin & Review, 21*(1), 78–84.

Oppermann, F., Jescheniak, J. D., & Schriefers, H. (2008). Conceptual coherence affects phonological activation of context objects during object naming. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*(3), 587–601.

Ouyang, G., Sommer, W., Zhou, C. S., Aristei, S., Pinkpank, T., & Rahman, R. A. (2016). Articulation artifacts during overt language production in event-related brain potentials: Description and correction. *Brain Topography, 29*(6), 791–813.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology-Learning Memory and Cognition, 24*(3), 539–557.

Python, G., Fargier, R., & Laganaro, M. (2018). ERP evidence of distinct processes underlying semantic facilitation and interference in word production. *Cortex, 99*, 1–12.

Qu, Q. Q., Damian, M. F. & Kazanina, N. (2012). Sound-sized segments are significant for Mandarin speakers. *Proceedings of the National Academy of Sciences of the United States of America, 109*(35), 14265–14270.

Qu, Q. Q., Damian, M. F., Zhang, Q. F., & Zhu, X. B. (2011). Phonology contributes to writing: Evidence from written word production in a nonalphabetic script. *Psychological Science, 22*(9), 1107–1112.

Rahman, R. A., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes, 24*(5), 713–734.

Rahman, R. A., & Melinger, A. (2019). Semantic processing during language production: An update of the swinging lexical network. *Language, Cognition and Neuroscience, 34*(9), 1176–1192.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1-3), 107–142.

Roelofs, A. (1997). The weaver model of word-form encoding in speech production. *Cognition, 64*(3), 249–284.

Roelofs, A. (2008). Tracing attention and the activation flow in spoken word planning using eye movements. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*(2), 353–368.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(3), 686–698.

Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(4), 896–918.

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*(1), 58–76.

Wang, C. & Zhang, Q. F. (2015). Phonological codes constrain output of orthographic codes via sublexical and lexical routes in Chinese written production. *PLoS ONE, 10*(4), e0124470.

Zhang, Q. F., Chen, H.-C, Weekes, B. S., & Yang, Y. F. (2009). Independent effects of orthographic and phonological facilitation on spoken word production in Mandarin. *Language and Speech, 52*, 113–126.

Zhang, Q. F., & Damian, M. F. (2010). Impact of phonology on the generation of handwritten responses: Evidence from picture-word interference tasks. *Memory and Cognition, 38*(4), 519–528.

Zhang, Q. F., & Damian, M. F. (2019). Syllables constitute proximate units for Mandarin speakers: Electrophysiological evidence from a masked priming task. *Psychophysiology, 56*(4), e13317.

Zhang, Q. F., & Weekes, B. S. (2009). Orthographic facilitation effects on spoken word production: Evidence from Chinese. *Language and Cognitive Processes, 24*(7-8), 1082–1096.

Zhang, Q. F., & Yang, Y. F. (2003). The determiners of picture-naming latency. *Acta Psychologica Sinica, 35*(4), 447–454.

[张清芳, 杨玉芳. (2003). 影响图画命名时间的因素. *心理学报, 35*(4), 447–454.]

Zhang, Q. F., & Yang, Y. F. (2006). The interaction of lexical selection and phonological encoding in Chinese word production. *Acta Psychologica Sinica, 38*(4), 480–488.

[张清芳, 杨玉芳. (2006). 汉语词汇产生中词汇选择和音韵

编码之间的交互作用. *心理学报, 38*(4), 480–488.]
Zhang, Q. F., & Zhu, X. B. (2016). It is not necessary to retrieve the phonological nodes of context objects for Chinese speakers. *Frontiers in Psychology, 7*, 1161.
Zhang, Q. F., Zhu, X. B., & Damian, M. F. (2018). Phonological activation of category coordinates in spoken word production: Evidence for cascaded processing in English but not in Mandarin. *Applied Psycholinguistics, 39*(5), 835–860.
Zhao, H. R., La Heij, W., & Schiller, N. O. (2012). Orthographic and phonological facilitation in speech production: New evidence from picture naming in Chinese. *Acta Psychologica, 139*(2), 272–280.
Zhou, X. L., & Marslen-Wilson, W. (1999). Sublexcial processing in reading Chinese. In J. Wang, A. Inhoff, & H-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 37–63). Hillsdale, NJ: Erlbaum.
Zhou, X. L., & Marslen-Wilson, W. (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1245–1265.
Zhu, X. B., Damian, M. F., & Zhang, Q. F. (2015). Seriality of semantic and phonological processes during overt speech in Mandarin as revealed by event-related brain potentials. *Brain and Language, 144*, 16–25.
Zhu, X. B., Zhang, Q. F., & Damian, M. F. (2016). Additivity of semantic and phonological effects: Evidence from speech production in Mandarin. *The Quarterly Journal of Experimental Psychology, 69*(11), 2285–2304.
Zhuang, J., & Zhou, X. L. (2003). The interaction between semantics and phonology in the speech production of Chinese. *Acta Psychologica Sinica, 35*(3), 300–308.
[庄捷, 周晓林. (2003). 汉语词汇产生中语义、语音层次之间的交互作用. *心理学报, 35*(3), 300–308.]

The multiple phonological activation in Chinese spoken word production: An ERP study in a word translation task

ZHANG Qingfang[1], QIAN Zongyu[1], ZHU Xuebing[2]

([1] *Department of Psychology, Renmin University of China, Beijing* 100872, *China*)

([2] *Shanghai International Studies University, Shanghai* 201620, *China*)

Abstract

A debatable issue between serial discrete models and interactive models is whether non-target lemmas activate their phonological words in spoken word production. Serial discrete models assume that only target lemma activates its corresponding phonological node to articulation, whereas interactive models assume that the semantic and phonological nodes linked to multiple candidates are co-activated during the retrieval of target word. Multiple phonological activation has been supported by evidences from alphabetic languages, but it remains unknown whether this finding can be generalized to non-alphabetic languages. Therefore, the current study aimed to investigate whether the not-to-be named pictures activate their phonological nodes in Chinese spoken word production.

Using electrophysiological measures, the present study employed a word translation task in native Chinese speakers with a high level of English proficiency. Thirty-two participants (13 males, average 22.94 years) were presented with an English probe word and a context picture (semantically related or unrelated, phonologically related or unrelated to target word) simultaneously. Eighty-six English probe words from CELEX database and forty-three black and white line pictures from a standardized picture database in Chinese were chosen as stimuli. Participants were asked to translate English probe words into Chinese as accurately and quickly as possible while ignoring context pictures presented simultaneously.

Behavioral results showed a typical semantic facilitation effect, with faster translation latencies in the semantically related condition than in the semantically unrelated condition. More importantly, phonological overlap, which generally elicits priming in Indo-European languages, resulted in a null finding for Chinese production. Electrophysiological results revealed that semantic relatedness induced significant effects of ERPs after stimuli presentation: a widely distributed positivity in the 400- to 600-ms interval, while marginally significant effects were observed for phonological relatedness in the time interval of 600~700 ms in the right middle region. Furthermore, a negative correlation between the difference of translation latencies (semantically related minus semantically unrelated) and the difference of mean amplitudes (semantically related minus semantically unrelated) approached significance in the 400~600 ms time window in the middle posterior region,

suggesting that more positive mean amplitudes were associated with shorter translation latencies.

Although speakers present a weak but reliable neural activation, we suggest that phonological overlap between context pictures and target words had no impact on the translation processing in behavioral. That is, the non-target lemma did not activate their phonological node, and multiple phonological activation was absent in Chinese spoken production. Meanwhile, the semantic information of context pictures was indeed activated, and according to the temporal course of word translation, the time window of 400~600 ms was estimated for conceptual preparation when Chinese-English bilinguals completed a word translation task, although this activation was not transmitted from semantic level to phonological level. Overall, the present findings support a serial discrete model rather than an interactive model in Chinese spoken word production.

Key words spoken word production, multiple phonological activation, serial discrete models, interactive models